УДК 004.932.2

05.00.00 Технические науки

СРАВНИТЕЛЬНЫЙ АНАЛИЗ НЕКОТОРЫХ АЛГОРИТМОВ РОЕВОГО ИНТЕЛЛЕКТА ПРИ ОБНАРУЖЕНИИ СЕТЕВЫХ АТАК НЕЙРОСЕТЕВЫМИ МЕТОДАМИ

Частикова Вера Аркадьевна
к.т.н., доцент
РИНЦ SPIN-код: 4525-0290
chastikova_va@mail.ru

Малыхина Мария Петровна
к.т.н., профессор
РИНЦ SPIN-код: 3848-6630
malpema@mail.ru

Жерлицын Сергей Анатольевич
студент
kpytooooo@gmail.com

Воля Яна Игоревна
студент
volya_y@mail.ru
Кубанский государственный технологический университет, Краснодар, Россия

В статье рассматривается проблема выявления сетевой атаки с целью последующего применения мер по обеспечению информационной безопасности. Для решения данной задачи проведено исследование эффективности работы нейронной сети с использованием в качестве алгоритмов обучения ряда метаэвристических методов, таких как, генетический алгоритм, алгоритм серых волков и алгоритм светлячков. Описаны механизмы указанных алгоритмов и принципы их работы. Для определения наличия сетевой атаки выбрана архитектура многослойного персептрона с сигмоидальной функцией активации. Исследованы различные конфигурации нейронной сети с целью определения оптимальной, имеющей соотношение слоев и нейронов на каждом слое, которое позволит получить минимальную ошибку. Обучение проводилось путем минимизации квадрата отклонения полученного выхода сети от эталонного с помощью заявленных алгоритмов. Генетический алгоритм требует тщательного подбора параметров при любом изменении структуры нейронной сети, а также значительно уступает в быстродействии алгоритмам светлячков и серых волков. Наибольшую точность показал метод серых волков, однако с увеличением количества скрытых слоев сети, его эффективность снижается. Большую устойчивость точности полученных результатов к изменению структуры нейронной сети показал алгоритм светлячков: он незначительно уступает по эффективности алгоритму волков, однако является более универсальным

UDC 004.932.2

Technical sciences

COMPARATIVE ANALYSIS OF SOME SWARM INTELLIGENCE ALGORITHMS WITH DETECTION OF NETWORK ATTACKS USING NEURAL NETWORK METHODS

Chastikova Vera Arkadyevna
Cand.Tech.Sci., associate professor
RSCI SPIN-code: 4525-0290
chastikova_va@mail.ru

Malykhina Maria Petrovna
Cand.Tech.Sci., professor
RSCI SPIN-code: 3848-6630
malpema@mail.ru

Zherlitsyn Sergey Anatolyevich
student
kpytooooo@gmail.com

Volya Yana Igorevna
student
volya_y@mail.ru
Kuban State Technological University, Krasnodar, Russia

This article is devoted to the problem of network attacks recognition, which is essential for providing network security. A research of neural network efficiency has been held. Such metaeuristic algorithms as genetic algorithm, gray wolf algorithm and firefly algorithm have been applied for the neural network learning. The algorithms' fundamentals have been described. Multilayer perseptrone with sigmoid activation function has been selected for the task of network attack presence check. Various configurations of the neural network have been tested in order to find the optimal number of layers and neurons per layer, which ensure the least error. Learning has been performed by minimization of the average squared error between the network's output and its target value with the help of the listed algorithms. Genetic algorithm requires accurate parameter picking in case of any network's architecture alteration. Moreover, it is not as fast as firefly and gray wolf algorithms. Gray wolf algorithm appears to be the most effective one. However, it loses its efficiency if the number of layers is increased. Firefly algorithm proves to be the most universal one. Although it is less effective than gray wolf algorithm, it provides the most exact output even if the network's structure is changed

Ключевые слова: СЕТЕВАЯ АТАКА, НЕЙРОННАЯ СЕТЬ, ГЕНЕТИЧЕСКИЙ АЛГОРИТМ, АЛГОРИТМ СЕРЫХ ВОЛКОВ, АЛГОРИТМ СВЕТЛЯЧКОВ

Keywords: NETWORK ATTACK, NEURAL NETWORK, GENETIC ALGORITHM, GRAY WOLF ALGORITHM, FIREFLY ALGORITHM

Doi: 10.21515/1990-4665-129-009

Важным аспектом существования информационного общества является стабильный доступ к базам данных и знаний в любой момент времени. Его обеспечение может быть скомпрометировано отсутствием или ненадежностью средств фильтрации трафика, то есть обнаружения сетевых атак. Целью таких атак и является, прежде всего, нарушение работы системы или затруднение доступа к информации, что может привести к колоссальным убыткам пострадавших лиц и компаний.

Данная работа посвящена исследованию и разработке средств защиты, отличающих вредоносные подключения от нормальных с целью последующего принятия мер по обеспечению безопасности системы.

Для обнаружения сетевых атак используются подходы на основе статистических методов, экспертных систем, нейронных сетей, а также системы, использующие эффективные виды гибридизации различных методов, обеспечивающие усиление качественных показателей друг друга. В статье рассматривается подход к решению данной задачи на основе гибридизации механизмов нейронных сетей и алгоритмов роевого интеллекта.

В ходе исследования были проанализированы различные архитектуры нейронных сетей. Из-за способности симулировать множественные условные взаимосвязи для решения поставленной задачи была выбрана архитектура многослойного персептрона (рис. 1). Подобная нейронная сеть принимает на вход вектор входных значений - параметры сетевого трафика. Каждый нейрон связан со всеми элементами последующего слоя, нейроны имеют сигмоидальную функцию активации.

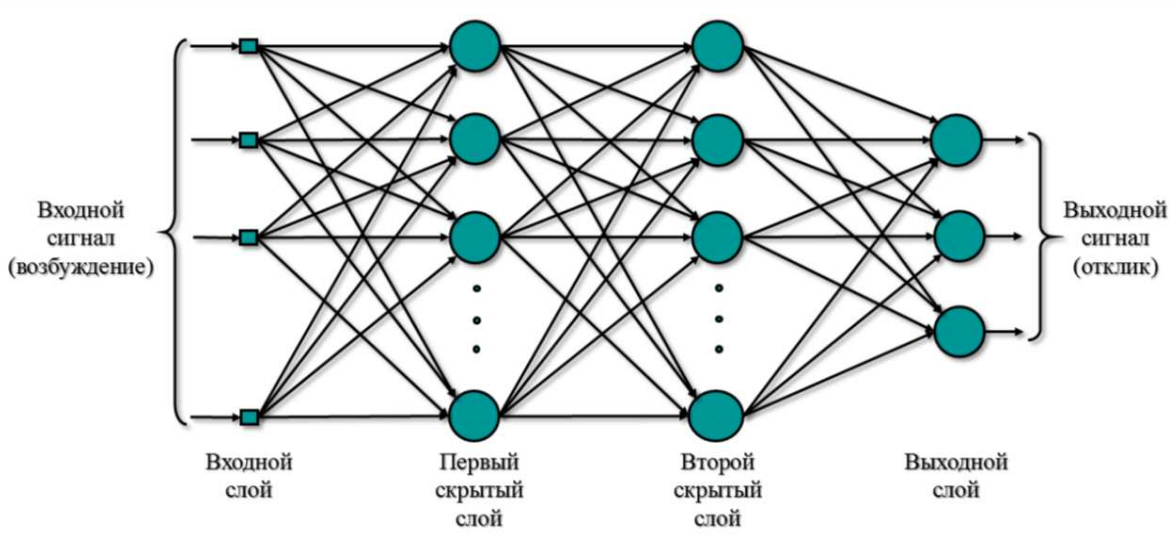

Рисунок 1 – Структура многослойного персептрона

Для обучения нейронной сети использовалась база данных сетевых атак, находящаяся в открытом доступе. На ее основе были отобраны 28 параметров сетевого трафика, позволяющих определить наличие вредоносной активности. Данные были конвертированы из текстового формата в цифровой и нормализованы для корректной работы нейронной сети.

Для обучения сети были применены: генетический алгоритм, алгоритм серых волков и алгоритм светлячков.

1. Генетический алгоритм

Генетический алгоритм основан на механизмах естественной эволюции в природе – наследовании и естественном отборе [2,3,6]. В работе используются операторы генетического алгоритма со следующими настройками:

- Наследование. Случайным образом выбираются 2 родительские пары хромосом на основе механизма рулетки и посредством оператора кроссинговера формируется дочерняя особь. При этом дочерние особи замещают наименее приспособленных особей предыдущего поколения.

- Одноточечный кроссинговер. Дочерняя хромосома условно разделена на 2 части, первая часть унаследована от первого родителя, вторая - от второго. Точка разделения выбирается случайным образом.
- Рулетка - механизм отбора родительских пар. Колесо рулетки включает по 1 сектору для каждой особи. Размер сектора пропорционален значению целевой функции особи, а значит вероятности быть отобранным для создания нового поколения.
- Мутации. Важный механизм, который позволяет алгоритму не зацикливаться на каком-либо локальном экстремуме. С вероятностью в 2% случайно выбранный в особи ген может быть заменен на случайное допустимое число.

В качестве фитнесс-функции использовано среднеквадратичное отклонение ожидаемых результатов, полученных на выходе нейронной сети с рассчитанными рассматриваемым алгоритмом весами, от значений, хранящихся в базе. Первоначальные значения особей инициализируются по следующей формуле:

$$x_i = 0{,}7 * a * (N - Inp - Out)^{\frac{1}{Inp}},$$

где x_i - *i*-ый элемент хромосомы (ген), a - случайно заданное целое число из диапазона [1, 135), N - количество нейронов в сети, Inp - размер массива, содержащего входные данные, Out - размер массива, содержащего выходные данные.

Далее популяция особей подвергается отбору и ряду генетических операторов, после чего происходит перерасчет значений функции приспособленности. Результатом работы алгоритма является массив особей, который представляет собой набор весов нейронной сети,

позволяющей определить тип действующей сетевой атаки или ее отсутствие.

2. Метод серых волков

Алгоритм серых волков, как и алгоритм светлячков, являются представителями метаэвристических алгоритмов роевого интеллекта. Они появились относительно недавно, но уже показали свою высокую эффективность при решении ряда оптимизационных задач, в том числе, задачи глобальной оптимизации многоэкстремальных функций [5,8].

Алгоритм серых волков был создан по подобию механизма охоты серых волков в природе и направлен на оптимизацию функции в векторном пространстве. Каждый поисковый агент (боид) принадлежит одному из 4 иерархических типов: альфа, бета, дельта и омега. В процессе выполнения оптимизации волки выполняют 3 действия: поиск добычи, окружение добычи и её атака.

Социальная иерархия выявляет трёх лучших волков и присваивает им ранги альфа, бета и дельта, по мере удаления от искомой точки. Все остальные волки получают ранг омега, таким образом в стае реализуется разделение на классы.

Основную часть работы алгоритма составляет цикл шагов, в котором волки окружают свою цель. Реализация окружения добычи заключается в обновлении позиций волков по следующим формулам:

$$\bar{D} = |\bar{C} * \bar{X}_p(t) - \bar{X}(t)|$$

$$\bar{X}(t+1) = \bar{X}_p(t) - \bar{A} * \bar{D},$$

где t –итерация, на которой находится алгоритм в данный момент, A и C – вспомогательные векторы-коэффициенты, X_p – координаты положения одного из 3-х лучших волков, X – текущее местонахождение рассматриваемого волка.

Векторы A и C вычисляются по следующим формулам:

$$\bar{A} = 2\bar{a} * \bar{r}_1 - \bar{a}$$

$$\bar{C} = 2 * \bar{r}_2$$

При этом a с каждой итерацией линейно уменьшается от 2 до 0 на протяжение работы алгоритма, а r_1 и r_2 – случайные числа на отрезке [0;1].

Таким образом, агенты распознают местоположение жертвы по ее окружению.

$$\bar{D}_\alpha = |\bar{C}_1 * \bar{X}_\alpha - \bar{X}|$$

$$\bar{D}_\beta = |\bar{C}_2 * \bar{X}_\beta - \bar{X}|$$

$$\bar{D}_\delta = |\bar{C}_3 * \bar{X}_\delta - \bar{X}|$$

$$\bar{X}_1 = \bar{X}_\alpha - \bar{A}_1 * (\overline{D_\alpha})$$

$$\bar{X}_2 = \bar{X}_\beta - \bar{A}_2 * (\overline{D_\beta})$$

$$\bar{X}_3 = \bar{X}_\delta - \bar{A}_3 * (\overline{D_\delta})$$

$$\bar{X}(t+1) = \frac{\bar{X}_1 + \bar{X}_2 + \bar{X}_3}{3}$$

Данные уравнения переопределяют координаты каждого из омега-волков таким образом, что их новая позиция оказывается в случайном месте внутри круга, заданного позициями 3-х лучших волков. Подобным образом боиды приближаются к добыче.

Для решения поставленной задачи веса связей нейронов представлялись как координаты волков, в качестве фитнесс-функции использовалась та же, что и при работе генетического алгоритма. Экспериментальным путем были выбраны параметры метода, определяющие оптимальные результаты: 8 поисковых агентов, 1 итерация на обучающий пример, 200000 обучающих примеров.

3. Метод светлячков

Первоначально создается рой поисковых агентов, каждый из которых характеризуется позицией и интенсивностью свечения (яркостью). Позиция задается случайным образом из определенного промежутка. Яркость рассчитывается на основе позиции, при этом самый яркий и, соответственно, самый приспособленный светлячок находится ближе всего к решению поставленной задачи. Важными характеристиками алгоритма являются такие параметры, как γ - коэффициент поглощения света среды и β - привлекательность одного светлячка для другого, убывающая экспоненциально с увеличением расстояния. Экспериментально были получены следующие оптимальные значения: γ=1 и β=1. При выполнении алгоритма используются следующие формулы:

$$pos(i) = pos_0(i) + a * c + \beta * (pos(i) - pos(j))$$

$$\beta = \beta_0 * e^{-\gamma * r^2}$$

где - γ коэффициент поглощения света среды, r - расстояние между особями, $pos_0(i)$ - начальное положение i, $pos(i)$ - новое положение i, β-привлекательность j для i, β_0 - привлекательность светлячка при отсутствии расстояния, то есть, когда они находятся вплотную друг к другу, c - случайно выбранное число из промежутка [0; 1], a - свободный параметр рандомизации. При этом боид с меньшей интенсивностью свечения перемещается к боиду с большей интенсивностью.

Для обучения веса связей нейронов представлялись как позиции светлячков, в качестве фитнесс-функции использовалась функция, аналогичная фитнесс-функции в генетическом алгоритме. Экспериментальным путем были выбраны параметры алгоритма, дающие лучшие результаты: 25 поисковых агентов, 1 итерация на обучающий пример, 200000 обучающих примеров.

4. Результаты исследования

Для определения эффективности обучения были выбраны сети нескольких конфигураций, а именно: 28-1, 28-14-1, 28-28-1, 28-28-14-1, 28-28-28-1. Каждое число означает количество нейронов на соответствующем слое. Так, архитектура 28-14-1 содержит входной уровень из 28 нейронов, выходной уровень, содержащий 1 нейрон, а также скрытый слой, который состоит из 14 нейронов. В таблице 1 представлены результаты исследования эффективности обучения нейронных сетей разных структур каждым из заявленных алгоритмов. Тестирование проводилось по всей базе (более 3 миллионов примеров).

Таблица 1 - Эффективность обучения нейронной сети

	28-1	28-14-1	28-28-1	28-28-14-1	28-28-28-1
Генетический алгоритм	87,514%	60,444%	56,38%	33,786%	24,972%
Алгоритм серых волков	99,224%	99,309%	93,095%	93,125%	88,589%
Алгоритм светлячков	94,763%	94,19%	98,88%	98,923%	98,233%

Как показывают результаты исследований, генетический алгоритм требует тщательного подбора параметров при любом изменении структуры нейронной сети, а также значительно уступает в быстродействии алгоритмам светлячков и серых волков. Наибольшую точность показал алгоритм серых волков (конфигурация 28-14-1), однако с увеличением количества скрытых слоев сети, его эффективность снижается. Большую устойчивость эффективности к изменению конфигурации нейронной сети показал алгоритм светлячков: он незначительно уступает по эффективности алгоритму волков, однако является более универсальным.

Литература

1. Частикова В.А., Власов К.А., Картамышев Д.А. Обнаружение DDoS-атак на основе нейронных сетей с применением метода роя частиц в качестве алгоритма обучения // Фундаментальные исследования. 2014. № 8-4. С. 829-832.

2. Симанков В.С., Частикова В.А. Генетический поиск решений в экспертных системах. - Краснодар, 2008.

3. Частикова В.А. Идентификация механизмов реализации операторов генетического алгоритма в экспертных системах продукционного типа //Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. 2012. № 75. С. 308-320.

4. Частикова В.А., Картамышев Д.А., Власов К.А. Нейросетевой метод защиты информации от DDoS-атак // Современные проблемы науки и образования. 2015. № 1-1. С. 183.

5. Частикова В.А., Жерлицын С.А. Исследование алгоритма серых волков // Материалы IV Международной научно-практической конференции «Автоматизированные информационные и электроэнергетические системы». – Краснодар, 9-11 сентября 2016.

6. Малыхина М.П., Частикова В.А., Власов К.А. Исследование эффективности работы модифицированного генетического алгоритма в задачах комбинаторики // Современные проблемы науки и образования. 2013. № 3. С. 32.

7. Частикова В.А., Власов К.А. Разработка и сравнительный анализ эвристических алгоритмов для поиска наименьшего гамильтонова цикла в полном графе // Фундаментальные исследования. 2013. № 10-1. С. 63-67.

8. Частикова В.А., Воля Я.И. Анализ эффективности работы алгоритма светлячков в задачах глобальной оптимизации // Научные труды Кубанского государственного технологического университета. 2016. № 15. С. 105-111.

9. Белов Д.Л., Антипова О.Ю., Частикова В.А. Методы решения задач с конфликтными ситуациями в системах принятия решений // Труды Кубанского государственного технологического университета. 2000. Т. 7. № 1. С. 153-159.

10. Частиков А.П., Белов Д.Л. Структура регенеративной экспертной системы // Материалы VII Международной научно-практической конференции «Инновационные процессы в высшей школе». 2001. С. 107-108.

11. Малыхина М.П., Бегман Ю.В. Нейросетевая экспертная система на основе прецедентов для решения проблем обслуживания абонентов сотовой сети // Известия высших учебных заведений. Северо-Кавказский регион. Серия: Технические науки. 2009. № 3. С. 6-9.

12. Малыхина М.П., Бегман Ю.В. Нейросетевые экспертные системы: обучение нейронной сети // Труды Кубанского государственного технологического университета. 2005. Т. 25. № 3. С. 93-94.

References

1. Chastikova V.A., Vlasov K.A., Kartamyshev D.A. Obnaruzhenie DDoS-atak na osnove nejronnyh setej s primeneniem metoda roja chastic v kachestve algoritma obuchenija // Fundamental'nye issledovanija. 2014. № 8-4. S. 829-832.

2. Simankov V.S., Chastikova V.A. Geneticheskij poisk reshenij v jekspertnyh sistemah. - Krasnodar, 2008.

3. Chastikova V.A. Identifikacija mehanizmov realizacii operatorov geneticheskogo algoritma v jekspertnyh sistemah produkcionnogo tipa //Politematicheskij setevoj

jelektronnyj nauchnyj zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta. 2012. № 75. S. 308-320.

4. Chastikova V.A., Kartamyshev D.A., Vlasov K.A. Nejrosetevoj metod zashhity informacii ot DDoS-atak // Sovremennye problemy nauki i obrazovanija. 2015. № 1-1. S. 183.

5. Chastikova V.A., Zherlicyn S.A. Issledovanie algoritma seryh volkov // Materialy IV Mezhdunarodnoj nauchno-prakticheskoj konferencii «Avtomatizirovannye informacionnye i jelektrojenergeticheskie sistemy». – Krasnodar, 9-11 sentjabrja 2016.

6. Malyhina M.P., Chastikova V.A., Vlasov K.A. Issledovanie jeffektivnosti raboty modificirovannogo geneticheskogo algoritma v zadachah kombinatoriki // Sovremennye problemy nauki i obrazovanija. 2013. № 3. S. 32.

7. Chastikova V.A., Vlasov K.A. Razrabotka i sravnitel'nyj analiz jevristicheskih algoritmov dlja poiska naimen'shego gamil'tonova cikla v polnom grafe // Fundamental'nye issledovanija. 2013. № 10-1. S. 63-67.

8. Chastikova V.A., Volja Ja.I. Analiz jeffektivnosti raboty algoritma svetljachkov v zadachah global'noj optimizacii // Nauchnye trudy Kubanskogo gosudarstvennogo tehnologicheskogo universiteta. 2016. № 15. S. 105-111.

9. Belov D.L., Antipova O.Ju., Chastikova V.A. Metody reshenija zadach s konfliktnymi situacijami v sistemah prinjatija reshenij // Trudy Kubanskogo gosudarstvennogo tehnologicheskogo universiteta. 2000. T. 7. № 1. S. 153-159.

10. Chastikov A.P., Belov D.L. Struktura regenerativnoj jekspertnoj sistemy // Materialy VII Mezhdunarodnoj nauchno-prakticheskoj konferencii «Innovacionnye processy v vysshej shkole». 2001. S. 107-108.

11. Malyhina M.P., Begman Ju.V. Nejrosetevaja jekspertnaja sistema na osnove precedentov dlja reshenija problem obsluzhivanija abonentov sotovoj seti // Izvestija vysshih uchebnyh zavedenij. Severo-Kavkazskij region. Serija: Tehnicheskie nauki. 2009. № 3. S. 6-9.

12. Malyhina M.P., Begman Ju.V. Nejrosetevye jekspertnye sistemy: obuchenie nejronnoj seti // Trudy Kubanskogo gosudarstvennogo tehnologicheskogo universiteta. 2005. T. 25. № 3. S. 93-94.